How Minds Work

Memories and Learning

Stan Franklin

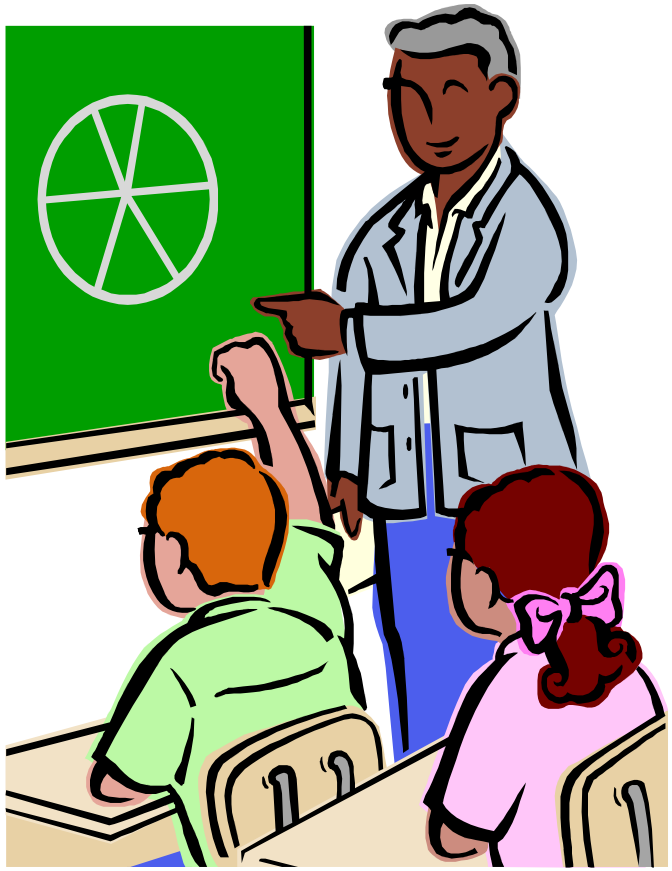
Computer Science Division &

Institute for Intelligent Systems

The University of Memphis

Human Learning

- More or less
  - Continual
  - Quick
  - Efficient
  - Accurate
  - Robust
  - Flexible
  - Effortless



Problems with Machine Learning

- Requires large, accurate training sets
- Little awareness of what's known or not known
- Integrates new knowledge poorly into old
- Learns one task at a time
- Little transfer of learned knowledge to new tasks
- Poor at learning from human teachers

For more human-like learning in machines

Design autonomous software agents and mobile robots using principles from human learning

Some Principles of Human Learning

- There's no learning from scratch
- We learn what we attend to
- Learning is a trial and error process
- Much memory is associative and content addressable

Lessons for Machine Learning

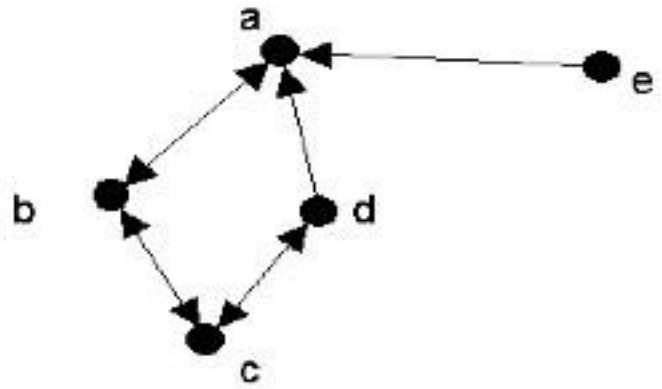
- Build in
  - Primitive feature detectors
  - Preferences for learnings
  - Attention mechanism
  - Base level activation
  - Inverse sigmoidal decay
- Make memory
  - Associative
  - Content addressable

Types of Human Learning Requiring Distinct Mechanisms

- Perceptual learning
  - Identify, classify, situate
- Episodic learning
  - What, where, when
- Procedural learning
  - New skills
  - Improve skills
  - Automatize
- Attentional learning
  - To what to attend

Examples of Learning Mechanisms

- Perceptual learning via a Slipnet
- Episodic learning via Sparse Distributed Memory
- Procedural learning via a Schema Net



A mechanism for perceptual learning

- Semantic net with activation passing
- Nodes represent features, individuals, categories, ideas
- Links represent relations, some excitatory, some inhibitory
- Total activation = current + base-level
- Percept composed of nodes over threshold
- Learning modifies base-level activation

A mechanism for episodic learning

- Sparse distributed memory
  - Few hard locations, HUGE address space
  - Each hard location contributes to the encoding of many different events
  - Each event is encoded to many different hard locations
- Associative and content addressable
- Psychological properties
  - Knows when it doesn't know
  - Tip of the tongue phenomena

A Mechanism for Procedural Learning

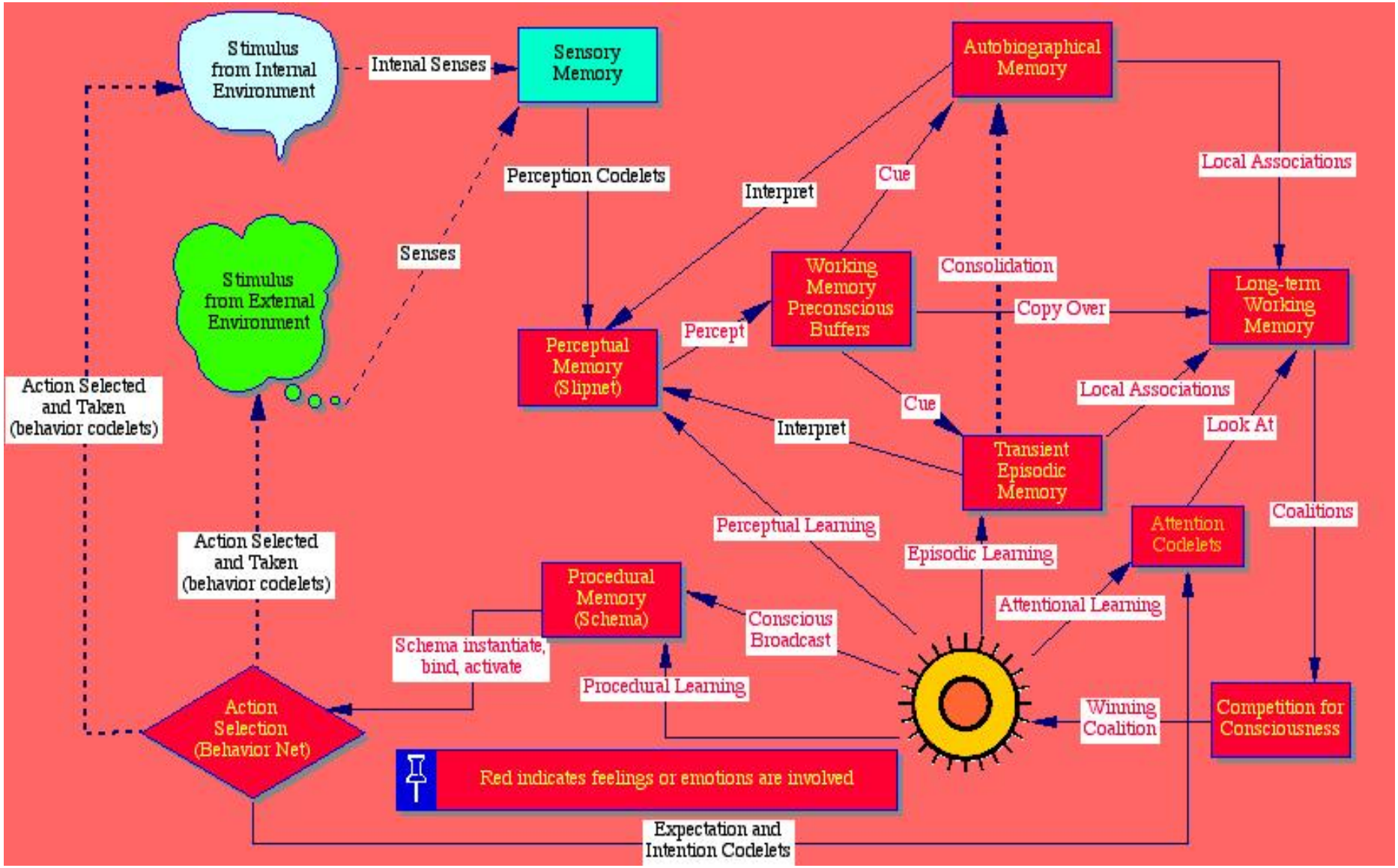
- Procedural learning via a schema net
- Activation = current + base-level
- Procedural learning
  - Reinforces base-level activation
  - Forms new schemas
- Schemas instantiate, activate, bind
- Incremental learning produces new behavior streams

Cognitive Cycles

- Every autonomous agents operates in the world by frequent, probably cascading, sense-process-act cycles
- Learning takes place during each cycle
- Learning is a function of attention and of arousal level.
- Feelings and emotions modulate learning

Human Cognitive Cycle Processing

- Hypothesis— Human cognitive processing is via a continuing sequence of Cognitive Cycles
- Duration— Each cognitive cycle takes roughly 200 ms
- Cascading— Several cycles may have parts running simultaneously in parallel
- Seriality— Consciousness maintains serial order and the illusion of continuity
- Start— Cycle may start with action selection instead of perception


Stimulus from Internal Environment
Intenal Senses
Sensory Memory
Autobiographical Memory
Local Associations
Perception Codelets
Interpret
Cue
Stimulus from External Environment
Senses
Working Memory Preconscious Buffers
Consolidation
Long-term Working Memory
Copy Over
Percept
Perceptual Memory (Slipnet)
Action Selected and Taken (behavior codelets)
Local Associations
Cue
Look At
Interpret
Transient Episodic Memory
Perceptual Learning
Episodic Learning
Attention Codelets
Coalitions
Action Selected and Taken (behavior codelets)
Procedural Memory (Schema)
Attentional Learning
Conscious Broadcast
Schema instantiate, bind, activate
Procedural Learning
Action Selection (Behavior Net)
Winning Coalition
Competition for Consciousness
Red indicates feelings or emotions are involved
Expectation and Intention Codelets

Perceptual Associative Memory

- Ability to interpret incoming stimuli by
  - recognizing individuals
  - categorizing them
  - noting situations
- Ubiquitous among animal species
- Animals of all sorts can identify food sources, potential mates, potential predators, etc.

Examples of PAM

- Pigeons – can categorize photos using such concepts as tree, fish, and human
- Honey bees – can identify letters independently of size, color, position or font
- African Grey Parrot (Alex) – can identify such features as size, number, color, and material of (sets of) objects never before seen

Perceptual Learning Premises

- (Almost?) ubiquitous among animals
- Evolutionarily older than semantic memory
- Distinct neural mechanisms
- Conscious awareness sufficient
- Facilitated by feelings and emotions
- Decays by an inverse sigmoid function
- First step in a sense-cognize-act cycle

Sensation

- Sensory receptors are directed by action
  - Saccades of the eyes
  - Sniff
  - Turing of an ear
  - Sending of an echolocation signal
- The environment impinges on receptors

Perception

- Filters sensory input based on expectation
- Simultaneously attaches meaning to it
- Identifies individuals, categories, situations and feelings
- Produces a percept including individuals, categories, relations, ideas, and some interpreted sensory data, i.e, qualia

Perception in IDA's Cycle

- Input arrives through senses
- Perception codelets find features and activate appropriate nodes in the slipnet
- Activation passes from node to node until the slipnet stabilizes
- Streams from different senses converge & bits of meaning combine into larger chunks
- Nodes over threshold form the percept
- Sensory stimuli received & interpreted, producing initial meaning

Each Node a Feature Detector

- Primitive feature detector – direct connection to receptive field
- Higher level feature detectors combine lower level feature detectors
- Object nodes detect features of object
- Category nodes detect members as features, as well as other features

Perceptual Learning

- Using the broadcast contents of consciousness
- Strengthen (or weaken) existing objects, categories, ideas, relations, feelings, etc.
- Create new objects, categories, situations

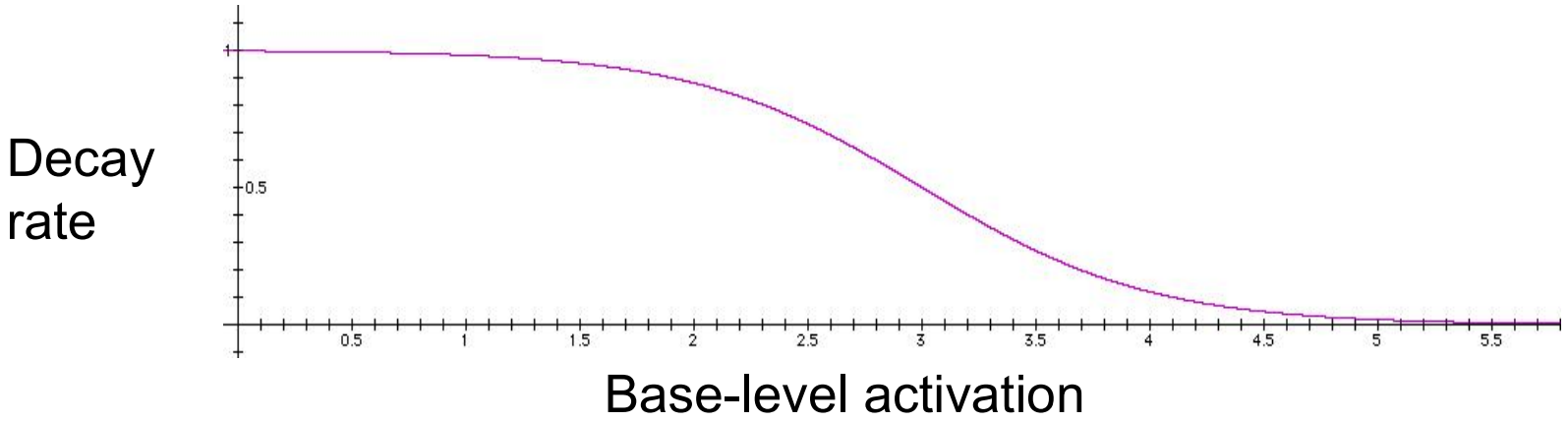
Contents of Consciousness

- Slipnet nodes are perceptual symbols
- Uniform representation throughout the IDA model
- Slipnet nodes comprise conscious contents

Modifying Existing Nodes

- Current activation
  - from other nodes
  - starting with primitive feature detectors
  - rapid decay
- Base-level activation
  - Inverse sigmoidal decay
- Total node activation a function of current and base level activations
- Perceptual learning modifies base-level activation of each node in conscious contents

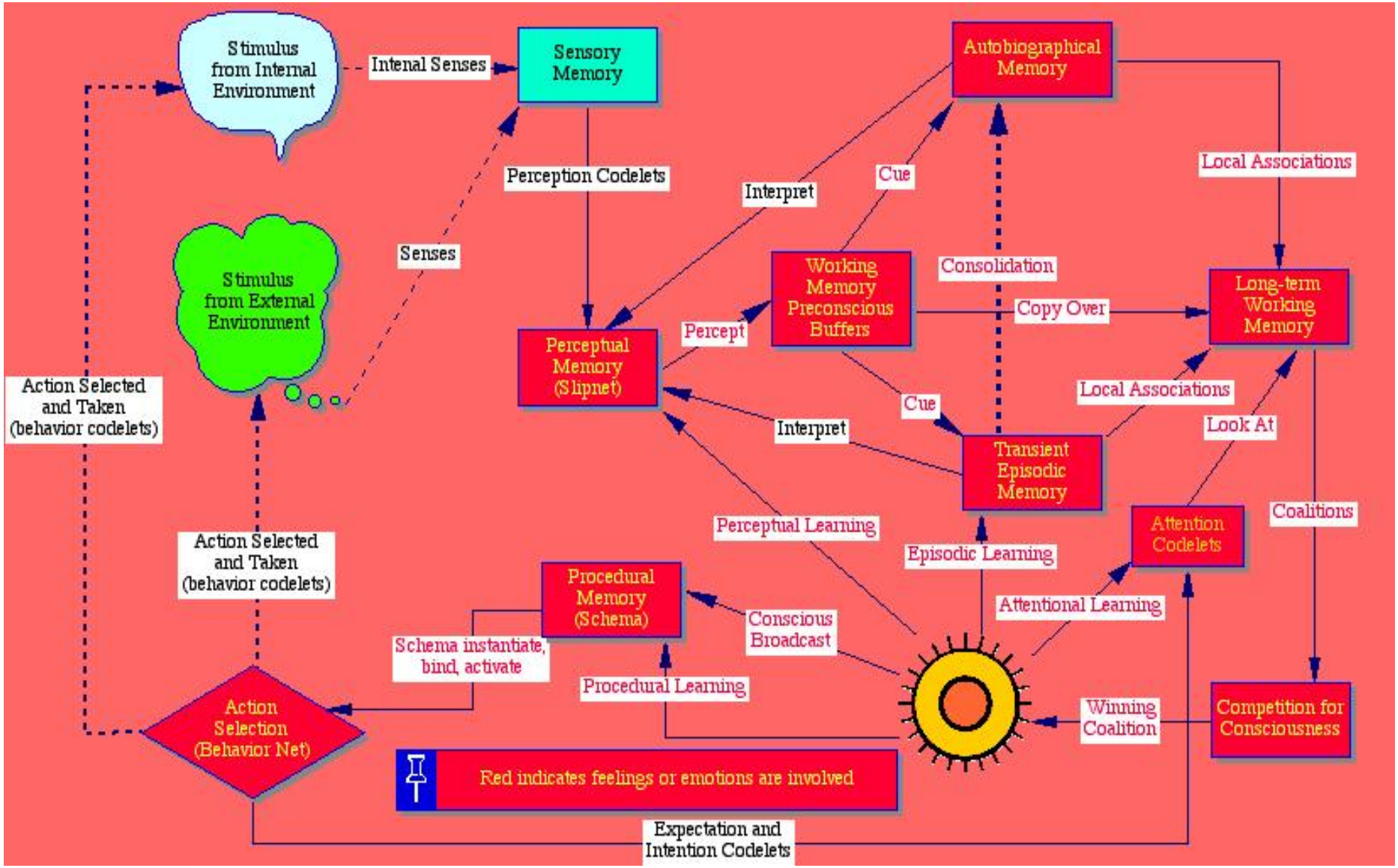
Inverse Sigmoidal Decay



- Low base-level activation = rapid decay
- Saturated base-level actibation = almost no decay

Learning New Nodes

- Object nodes – from noting contiguity of motion of features
- Category nodes – from noting similarity of objects
- Requires specific attention codelets
- A generate & test procedure due to inverse sigmoidal decay


Stimulus from Internal Environment
Intenal Senses
Sensory Memory
Autobiographical Memory
Local Associations
Perception Codelets
Interpret
Cue
Stimulus from External Environment
Senses
Working Memory Preconscious Buffers
Consolidation
Long-term Working Memory
Copy Over
Percept
Perceptual Memory (Slipnet)
Action Selected and Taken (behavior codelets)
Local Associations
Cue
Look At
Interpret
Transient Episodic Memory
Perceptual Learning
Episodic Learning
Attention Codelets
Coalitions
Action Selected and Taken (behavior codelets)
Procedural Memory (Schema)
Attentional Learning
Conscious Broadcast
Schema instantiate, bind, activate
Procedural Learning
Action Selection (Behavior Net)
Winning Coalition
Competition for Consciousness
Red indicates feelings or emotions are involved
Expectation and Intention Codelets

Transient Episodic Memory

- Records what, when, how, feelings, actions
- Content addressable from partial cues
- Consciously noted events are encoded
- Modulated by feelings and emotions
  - More affect, stronger encoding
  - More affect, more often in consciousness
- Decays in hours or a day

Writing to TEM

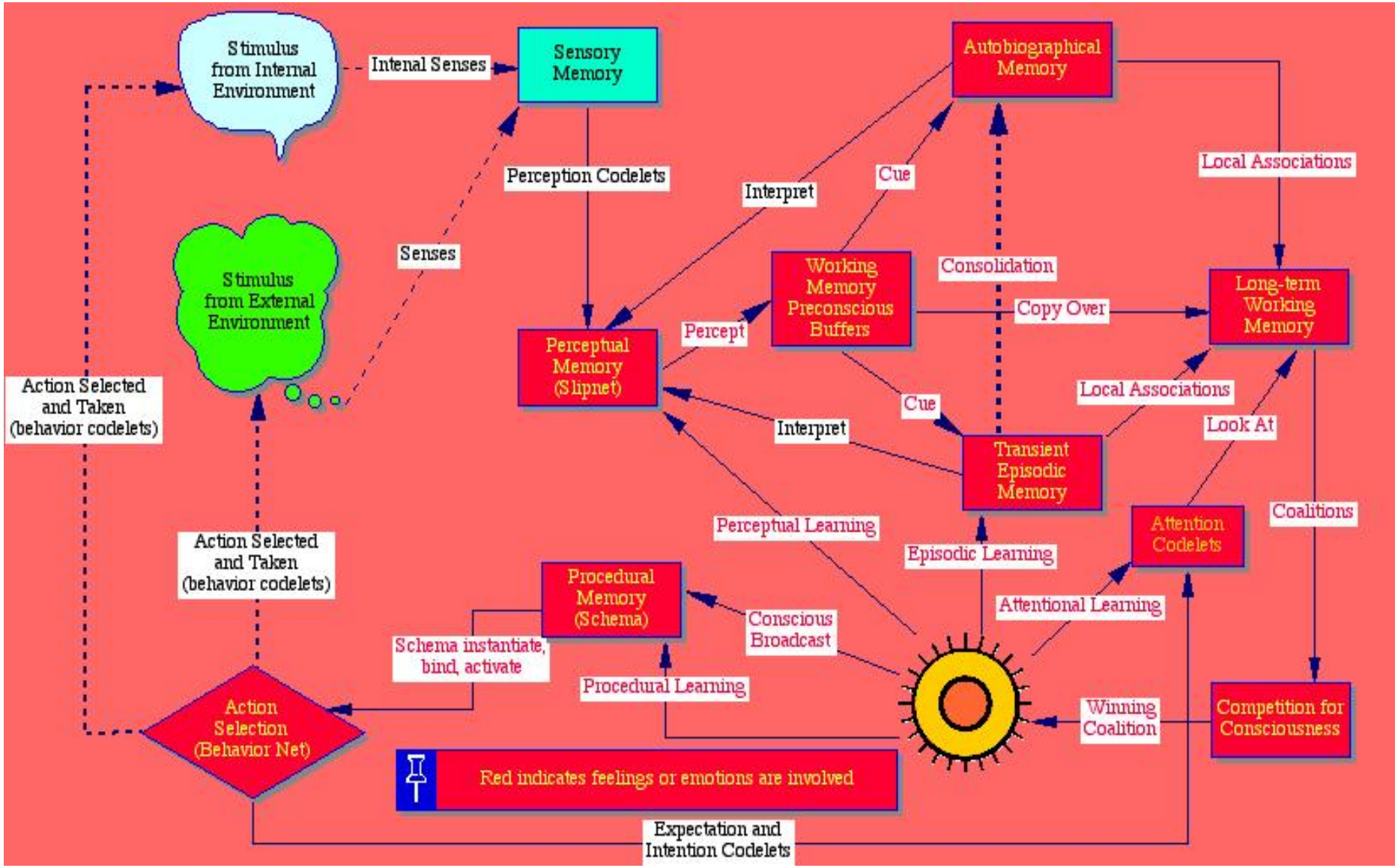
- Uses sparse distributed memory
- Perceptual symbols = universal representation
- Primitive feature detectors = subsets of dimensions
- Translate higher level feature detectors to primitive feature detectors for writing

Reading from TEM

- Contents of working memory form the cue
- Must be translated to primitive feature detectors
- Return from SDM routed through perception to produce perceptual symbols

Declarative Memory

- Autobiographical + semantic
- Read from just as TEM
- Write is consolidation from TEM
- Inverse sigmoidal decay
- Can decay very quickly
- Can last for decades


Stimulus from Internal Environment
Intenal Senses
Sensory Memory
Autobiographical Memory
Local Associations
Perception Codelets
Interpret
Cue
Stimulus from External Environment
Senses
Working Memory Preconscious Buffers
Consolidation
Long-term Working Memory
Copy Over
Percept
Perceptual Memory (Slipnet)
Action Selected and Taken (behavior codelets)
Local Associations
Cue
Look At
Interpret
Transient Episodic Memory
Perceptual Learning
Episodic Learning
Attention Codelets
Coalitions
Action Selected and Taken (behavior codelets)
Procedural Memory (Schema)
Attentional Learning
Conscious Broadcast
Schema instantiate, bind, activate
Procedural Learning
Action Selection (Behavior Net)
Winning Coalition
Competition for Consciousness
Red indicates feelings or emotions are involved
Expectation and Intention Codelets

Procedural Learning

- Learning new tasks - instructionist
- Reinforcing old tasks - selectionist
- Learning via consciousness
- Primitive effectors (motor neurons & muscles) not learned
- Very short term to very long term

Implementation

- Procedural learning via a schema net
- Schema (context, action, result)
- = behavior codelet in priming mode
- Primitive effector = empty schema (only action)
- Links from schema to derived schema

Activation of a Schema

- Modify as a result of conscious content
- Base-level activation – reinforce if action succeeded
- Current activation – depends on
  - Relevance of context to current situation
  - Relevance of result to current goals or feelings
- Total activation = base-level + current

Decay of Schema Activation

- Base-level activation – inverse sigmoidal
  - Low level schemas decay very rapidly
  - Saturated schemas decay hardly at all
  - Learning can be short or long term
- Current activation decays rapidly

Selectionist Learning

- Selectionist = reinforcement
  - Conscious broadcast says expectation met
  - Affect valence positive
  - Increase base-level activation in proportion to affect level
  - Valence negative or expectation not met
  - Decrease base-level activation in proportion to affect level

Instructionist Learning

- Behavior = goal context = schema with parallel compound action
- Merge two schema into a third
- Behavior stream = goal context hierarchy = schema with sequential compound action
- Merge two schema into a third

Attentional Learning

- Built in attention codelets for
  - Temporal sequence (= causality)
  - Similarity (= categorization)
  - Contiguity of motion (= object formation)
- Expectation codelet spawned with each executed behavior (codelet)
- Intention codelet produced with each volitional decision
- Others learned?

Email and Web Addresses

- Stan Franklin
  - franklin@memphis.edu
  - www.cs.memphis.edu/~franklin
- “Conscious” Software Research Group
  - www. csrg.memphis.edu/